Study Report:

SYSTEMS-BASED IDENTIFICATION OF PATIENTS WITH OSTEOPOROTIC VERTEBRAL FRACTURES

Jay Pandya, Markus Seibel, Lloyd Ridley, Kirtan Ganda

This study was supported by an unrestricted grant from HCF Research Foundation (Project ID: KGUS2015014) to The University of Sydney. The authors generated and analysed the data without input from the sponsor.

Authors' affiliation:

Dr Jay Pandya, Blacktown Hospital, Concord, NSW 2139 and The University of Sydney
Dr Lloyd Ridley, Radiology Departement, Concord Hospital, Concord NSW 2139
Prof Markus J Seibel, Dept of Endocrinology, Concord Hospital, and Bone Research Program, ANZAC Research Institute, Concord, NSW 2139
Dr Kirtan Ganda, Dept of Endocrinology, Concord Hospital and Bone Research Program, ANZAC Research Institute, Concord, NSW 2139

Correspondence to Dr Kirtan Ganda at Kirtan.ganda@sydney.edu.au

January 2017

Abstract / Summary

INTRODUCTION

The 'care gap' in osteoporosis management refers to the lack of assessment and appropriate management of osteoporosis despite patients suffering from a minimal trauma fracture. Secondary Fracture Prevention (SFP) programs have been established primarily to address this care gap. These programs identify, assess and manage those suffering from a recent minimal trauma fracture. However, SFP programs may miss patients with vertebral fractures, which confer a higher risk of re-fracture than non-vertebral fractures and are often asymptomatic. Therefore, it is vital to identify patient with vertebral fractures so that they can be appropriately managed. This could potentially be achieved through a systematised search of electronic radiology reports, aka Natural Language Processing (NLP) methods, however this had not been performed previously in the setting of a SFP program.

OBJECTIVES

This study aims to:

i) Develop an effective and reliable NLP method to systematically identify patients with radiographic vertebral fractures via searching free-text radiology reports.

ii) Determine the clinical utility of the report searches by attempting to invite eligible patients to attend the Concord Hospital SFP program.

METHODS

This was a retrospective observational study, which consisted of two phases:

1) The Development Phase whereby the output and positive predictive values for twelve search terms utilised in the identification of vertebral fractures from radiology reports was established, and

2) The Implementation Phase, which applied the three most effective search terms to identify patients with vertebral fractures who were then invited to attend the SFP program for further management as part of best clinical practice.

RESULTS

The Development phase revealed that the three search terms with the highest total output were ‘Loss of Height’, ‘Compression Fracture’ and ‘Crush Fracture’. During the Implementation phase, ‘Loss of Height’ was more effective compared to ‘Compression Fracture’ (OR 15.24, 95%CI 1.85-125.82, p=0.011) and ‘Crush Fracture’ (OR 15.00, 95%CI 1.22-183.63, p=0.034). The term ‘Compression Fracture’ was similar to ‘Crush Fracture’ in its ability to identify patients with vertebral fractures (OR 0.98, 95%CI 0.19-5.04, p=0.985).

During the Implementation Phase, a total of 126 radiology reports were identified over a 3-week period, of which 96 were unique patients. Amongst these patients, 69 (72%) were confirmed to have sustained a vertebral fracture. Of these 69 patients, 20 (29%) were invited to attend the Concord Hospital SFP program.

CONCLUSION

Simple NLP methods can be successfully utilised to identify patients with vertebral fractures via Electronic Medical Record (radiology reports) searches. The most effective terms were ‘Loss of Height’, ‘Compression Fracture’ and ‘Crush Fracture’. These NLP methods may be translated to other SFP programs to identify patients with vertebral fractures, thereby further narrowing the ‘care gap’ in osteoporosis management.

Acknowledgments

We would like to thank the HCF Research Foundation for agreeing to fund this project over the last year. We hope this study provides a valuable contribution in the field of bone health and osteoporosis.

We would also like to thank Dr. Mario D'Souza Research Statistician at Sydney Local Health District for his guidance relating to statististical analysis.

Table of Contents

List of Tables, Figures and Illustrations

Tables

Figures

List of Special Names or Abbreviations

EMRs = Electronic Medical Records

NLP = Natural Language Processing

SFP = Secondary Fracture Prevention

PPV = Positive Predictive Value

OR = Odds Ratio

TP = True Positive

FP = False Positive

Systems Based Identification of Patients with Osteoporotic Vertebral Fractures

INTRODUCTION / BACKGROUND

Osteoporosis is a disorder of low bone mass and micro-architectural deterioration of bone resulting in decreased mechanical strength leading to an increased susceptibility to fractures even after minimal trauma (1). These minimal trauma fractures, also referred to as osteoporotic or fragility fractures, represent a major public health problem, due to their high prevalence and associated clinical consequences (2).

In 2012, there were 140,822 minimal trauma fractures in Australia. By 2022, this figure is expected to increase to 183,105 (3). On statistical grounds, more than 50% of postmenopausal women, and 30% of men over the age of 60 years will suffer at least one osteoporotic fracture during their remaining lifetime (4, 5). Warr's et al, suggests that the prevalence of osteoporosis is likely to be significantly underestimated as the diagnosis is often made following symptomatic non-vertebral fractures (in both men and women). Consequently, of the 1.2 million Australians likely to have osteoporosis, most will not know that they have the disease. The total cost of osteoporosis in Australia in 2012 was $2.75 billion. This is expected to rise over the next ten years to $33.6 billion (3).

There is ample evidence that any fragility fracture predisposes to further fractures (6-9), significant morbidity and premature death (2, 10). However, spine or vertebral fractures portend a particularly high risk of re-fracture, with a five to ten-fold increased risk of further vertebral fracture, and a two-fold increased risk of hip fracture (9, 11). Importantly, vertebral fractures often occur with no obvious trauma and remain, at least initially, asymptomatic in 60% of cases. However, over time these fractures also cause chronic pain and disability, impact on respiratory reserve, reduce mobility and are associated with increased mortality (12).

For over two decades, we have known that the timely diagnosis and optimal treatment of osteoporosis prevents further fractures. Several safe and effective medications are available and osteoporosis guidelines around the world recommend long-term treatment for people who have sustained a minimal trauma fracture. Anti-osteoporosis treatments are particularly effective in preventing vertebral fractures, with relative risk reductions of up to 70% (13-16). However, the Australian and international literature provides ample proof that the majority of patients with vertebral osteoporotic fractures are neither assessed for osteoporosis, nor appropriately managed to prevent further fractures (17, 18). This is even more astounding in Australia, as effective pharmacotherapies are available on the Pharmaceutical Benefits Scheme for patients with proven vertebral (or other) minimal trauma fractures.

With the implementation of Electronic Medical Records (EMRs), new opportunities for quality improvement and research can be employed that were previously not feasible using what is termed Natural Language Processes (NLP) (19). NLP has been employed to interpret volumised textual data for over six decades, and it now permeates through daily life from internet search engines to telephone banking and e-mail filtering. Unfortunately, it has only recently gained ground in the field of medicine as a consequence of the adoption of EMRs in healthcare facilities (20). EMRs currently compose primarily unstructured textual data, hence the challenge lies in uncovering and extracting the necessary information from this rich data source (21). NLP is one method of data extraction. It works by recognizing pre-specified entities or targeted sequences of text, which is then used to extract sections of data that denote important information. Natural Language Processing techniques that are used to extract data from unstructured text into formal computer representations are a valuable tool for creating robust, scalable methods of mining data in EMRs, including radiology reports (22).

There have been a number of studies that have proven the benefits of NLP in healthcare. Recent studies that have employed NLP to extract data from radiology reports have stated that NLP has a significantly higher sensitivity when compared to other methods such as discharge coding systems (19, 21-23). One study quoted precision of 99.4% and recall of 99.1% with an overall accuracy of 96% (21).

It is critical to improve the identification of patients with osteoporotic vertebral fractures, which have a high prevalence, are clinically relevant and have highly effective treatments available. Moreover, these patients need to be directed towards services that investigate and manage their osteoporosis appropriately to reduce the risk of re-fracture. With the adoption of EMRs and the development of NLP, a new method now exists to identify these patients.

The established Secondary Fracture Prevention (SFP) program at Concord Repatriation General Hospital has a successful track record of identifying and managing patients with symptomatic osteoporotic fractures. Published long term follow-up data demonstrate significant reductions in re-fracture incidence of up to 80% (24). However, the overwhelming majority (>95%) of patients in the SFP program present with non-vertebral fractures because these fractures are symptomatic and therefore easier to identify. Thus, few vertebral fractures are detected through the SFP program, which are often asymptomatic. As a next step, therefore, the current study aims to develop effective and reliable methods and algorithms (NLP methods) to identify patients with radiographic vertebral fractures. Once identified, these patients will be referred to the SFP program at Concord Repatriation General Hospital for further assessment.

We hope that by utilizing modern techniques, we are able to provide a method by which to extend the scope of the current SFP program, thereby reducing the risk of further fractures in a population at immensely heightened risk of osteoporotic fracture. The long-term aim is to reduce health care cost and morbidity by preventing hospitalization due to re-fractures.

METHODS

Patient Population and Study Design

This was a retrospective, single centre observational study. It was primarily designed to apply Natural Language Processing in a clinical setting using pre-existing government departmental software. The primary research facility was the Department of Endocrinology and Metabolism, Concord Repatriation General Hospital in Sydney, Australia.

The study consisted of two phases i.e. 1) the Development phase and 2) the Implementation phase.

The Development phase involved the identification of key strings of texts or “pre-specified inputs” that we could use to mine data from our electronic radiology reports using the pre-existing search software, Centricity RIS-i 4.2 Plus. Through testing various terms by iterative processing, we aimed to produce effective and reliable sequences of texts that could identify patients with vertebral fractures. A vertebral fracture was defined as a 20% or greater reduction in height of the anterior or mid portion of a vertebral body relative to the posterior height of the body, or, a 20% or greater reduction in any of these heights compared to the vertebral body above or below the affected vertebral body.

A number of search terms were developed in consultation with a Radiologist to identify vertebral insufficiency fractures. Search terms included ‘Osteoporotic Fracture’, ‘Vertebral Compression Fracture’, ‘Wedge Fracture’, ‘Burst Fracture’, ‘Compression Fracture’, ‘Loss of Height’, ‘Collapse of Vertebral Body’, ‘Anterior Wedge Compression Fracture’, ‘Osteoporotic Compression Fracture’, ‘Insufficiency Fracture’, ‘Compression Deformities’, and ‘Crush Fracture’. Each search term was then individually run through the Centricity RIS-I 4.2 Plus radiology software using parameters to help limit the search to patients from Concord Hospital, X-Rays and CT Scans. All reports within the database were scanned. The software provided an output of patient identifying details (name and date of birth) as well as full text radiology reports. Theses details were then exported into a text file for further analysis of results. Two researchers reviewed the output for each search term spanning the last six months of 2015 to determine the presence or absence of vertebral fractures. Positive predictive values of the respective terms were then determined.

Phase two of the project (Implementation phase) involved the translation of NLP methods to the clinical environment. The three terms with the highest total output identified from the first phase were utilised. These three search terms were used to identify patients with vertebral fractures on radiology reports spanning a period of one week per month for three consecutive months. At the end of the selected week, a member of the research team ran the pre-specified text inputs (i.e. three search terms) using the Centricity RIS-I 1.4 application to identify patients with vertebral fractures. Radiology reports of all patients extracted were then manually reviewed to identify those patients with osteoporotic vertebral fractures. Male and female patients above the age of 50 years with osteoporotic vertebral fractures, were then invited to attend the SFP program for further assessment and management of their osteoporosis, as a part of best clinical practice.

Statistical Analysis

Statistical analysis was conducted using Microsoft Excel 2015 (V 15.13.3) and IBM SPSS (Version 20) software. In phase 1 of the project, positive predictive values of search terms were determined using the formula: positive predictive value = total true positive / (total true positive + false positive). The percentage of true positives for each search term was calculated using the formula: percentage true positive = true positive for individual search term / total true positives from all three search terms. In phase two, logistic regression analysis was used to compare the effectiveness of the three different search terms. We used the Fisher method of binomial logistic regression with fracture being the dependent variable; 1 is identified as a true positive and 0 false positive. Duplicates were excluded from the regression analyses as their rates were deemed too low amongst terms to significantly skew data. A p-value of < 0.05 was deemed significant. Demographic characteristics were described using medians and inter-quartile ranges for continuous variables and percentages for categorical variables.

Ethics

The study was conducted with ethics approval obtained from the Sydney Local Health District Human Research Ethics Committee - Concord Repatriation General Hospital (EC00118). Data collected was stored at Department of Endocrinology and Metabolism, Level 6 Concord Hospital Medical Centre, Concord Repatriation General Hospital. Electronic Excel spreadsheets of data were all password protected and located in a computer at the above-mentioned facility. Data was only accessible to research investigators and will be stored at the above-mentioned facility for up to 7 years, after which it will be permanently destroyed. Funding was received from HCF Australia to assist with costs of undertaking this study.

RESULTS

The Development Phase

A total of 909 radiology reports from the Concord Hospital radiology database were identified using twelve search terms during the development phase. After reviewing the radiology reports, there were a total of 689 (75.8%) true positive reports (TP) (i.e. reports that identified patients with a vertebral fracture) and 220 (24.2%) false positive (FP) reports (i.e. reports that incorrectly identified patients as sustaining a vertebral fracture). The results for each search term are detailed in Table 1.

The three search terms with the highest output of reports were ‘Loss of Height’, ‘Crush Fracture’ and ‘Compression Fracture’. These three search terms had a combined output of 761 radiology reports, 581 (76%) of which identified patients with a vertebral fracture. The search term ‘Loss of Height’ identified the largest number and percentage of vertebral fractures identified with the 3 search terms (Table 2).

Table 1. Number (%) of reports identified and positive predictive value for the identification of vertebral fractures with each search term during Development Phase

Search Term	Total Output (n, %)	Positive Predictive Value	True Positive	False Positive
Loss of Height	394 (43%)	0.93	367	27
Crush Fracture	205 (23%)	0.45	93	112
Compression Fracture	162 (18%)	0.75	121	41
Burst Fracture	71 (8%)	0.89	63	8
Insufficiency Fracture	26 (3%)	0.62	16	10
Vertebral Compression Fracture	22 (2%)	0.14	3	19
Wedge Fracture	21 (2%)	0.95	20	1
Osteoporotic Fracture	5 (0.6%)	1.00	5	0
Anterior Wedge Compression Fracture	2 (0.2%)	0.50	1	1
Compression Deformities	1 (0.1%)	0.00	0	1
Collapse of Vertebral Body	0	N/A	0	0
Osteoporotic Compression Fracture	0	N/A	0	0

Table 2. Percentage of reports with confirmed vertebral fractures amongst the three search terms with the highest output. *True Positive = vertebral fracture confirmed; **False Positive = no vertebral fracture identified.

Search Term	Output	True Positives*	False Positives**	Percentage of all True Positives (%)
Loss of Height	394	367	27	63.2%
Compression Fracture	162	121	41	20.8%
Crush Fracture	205	93	112	16.0%
TOTALS	**761**	**581**	**180**	**100%**

The Implementation Phase

During the Implementation phase, the three search terms with the highest output during the Development phase were used to identify relevant radiology reports. A total of 126 radiology reports spanning three weeks in three consecutive months were identified, of which 93 (73.8%) represented unique patients. 69 of these patients (74.2%) had a vertebral fracture identified on their radiology report and were thus deemed true positives (Fig 1). Conversely, there were 24 patients (25.8%) with false positive search results i.e. did not have a vertebral fracture identified after reviewing the report. The age and gender of these patients are described in Table 3. Patients with a vertebral fracture identified on their radiology report were older and more likely to be male, compared to those with no vertebral fracture identified.

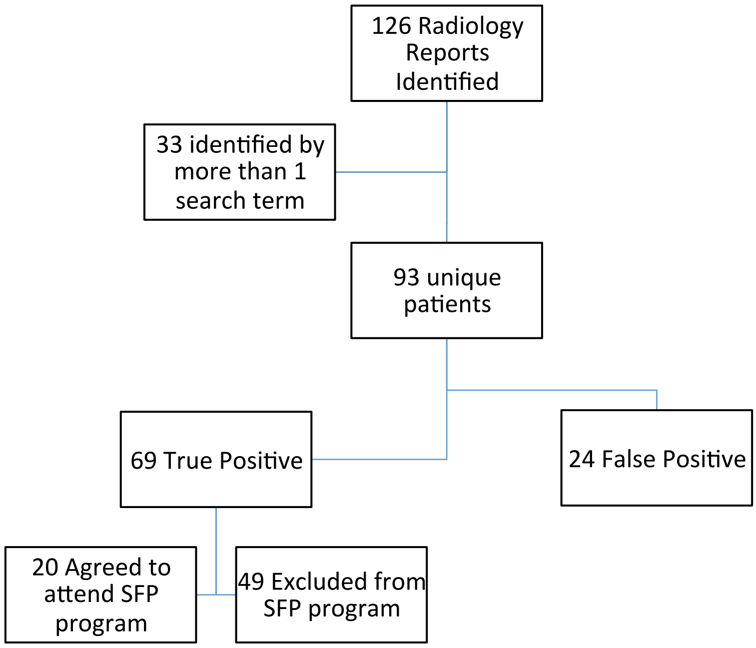


Figure 1. Flow diagram of radiology reports identified during Implementation phase

Table 3. Characteristics of patients identified during the Implementation phase

Group	Number	Median Age (years)	IQR	Female
All Patients	93	77.6	67.5 - 87.7	60%
Fracture	69	79.6	71.6 - 87.6	55%
No Fracture	24	62.7	50.5-75	65%

Of the 126 reports identified, only 27 (21.4%) were identified by more than one search term. Of these 27 reports, the terms ‘Loss of Height’ and ‘Compression Fracture’ had the highest overlap (Fig. 2), accounting for 66.7% (n=18) reports. There was minimal overlap between the other search terms (Fig. 2).

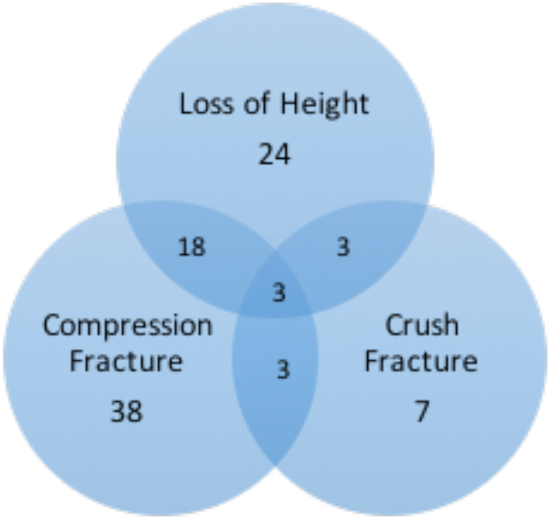


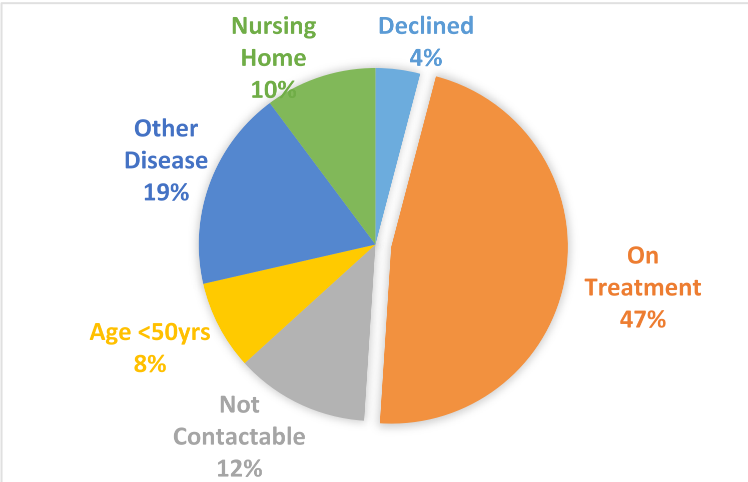
Figure 2. Venn Diagram demonstrating the number of reports of patients with vertebral fractures and their overlap, using the three most effective search terms.

Plain x-rays accounted for 54.8% (n=51), computed tomography 29.0% (n=27), magnetic resonance imaging 16.1% (n=15).

Logistic regression analysis comparing the three search terms' relative effectiveness in identifying vertebral fractures was completed after removing reports found by more than one search term. A total of 65 reports were analysed using logistic regression of which 45 (69.2%) had a fracture identified. The overall tests of model effects gave a global p-value of 0.039, suggesting all three terms were effective at identifying patients with fragility fractures. The term 'Loss of Height' was more effective in identifying fractures compared to 'Compression Fracture' (OR 15.24, 95%CI 1.85-125.82, p=0.011) and 'Crush Fracture' (OR 15.00, 95%CI 1.22-183.63, p=0.034). The term 'Compression Fracture' was similar to 'Crush Fracture' in its ability to identify patients with vertebral fractures (OR 0.98, 95%CI 0.19-5.04, p=0.985).

Of the 69 patients identified with vertebral fractures, a total of 20 patients (29.0%) agreed to attend the Concord Hospital SFP program. Amongst the remaining 49 patients, 2 declined, 23 were already receiving appropriate osteoporosis management, 6 were not contactable, 4 were < 50 years of age, 9 had a fracture secondary to another pathological process and 5 were residents of aged care facilities (Fig 3). From the 20 patients who agreed to attend the SFP program, the median age was 79.8yrs (IQR 72.5-87.1), which was similar to those were excluded (median age of 79.6yrs (IQR 71.4-87.8)). Females represented 45% of those who agreed to attend vs. 59% of those who were excluded from the SFP program.

Figure 3: Reasons patients with vertebral fractures did not attend the Secondary



DISCUSSION

Analysis of Results

This study is unique in that it is the first study using NLP methodology to identify patients with osteoporotic vertebral fractures as part of a SFP program. Using simple NLP methods, we successfully mined data from our existing radiology database to identify reports of patients with osteoporotic vertebral fractures. The Development phase identified the three key search terms, 'Loss of Height', 'Compression Fracture' and 'Crush Fracture', which represented 84% (n=761) of total radiology report output from all 12 terms. Of the 761 reports, 76% were confirmed vertebral fractures, which represents a high rate of vertebral fracture detection from the three search terms. The terms 'Compression Fracture' and 'Crush Fracture' had PPVs lower than many of the other 12 search terms, however the number of reports identifying vertebral fractures was significantly greater. Terms such as 'Osteoporotic Fracture' and 'Wedge Fracture' had high positive predictive values, however they had too low of an output to deem them clinically significant. Consequently, the terms 'Loss of Height', 'Compression Fracture' and 'Crush Fracture' were selected for use during the Implementation phase of the study. The main implication of our Development Phase is its benefit in the identification of the key search terms, which can subsequently be translated to any healthcare service with software enabling radiology report searches.

In the Implementation phase of the study approximately 42 records per week were extracted, which is significantly more than was initially expected. The terms overall performed well in identifying radiology reports of vertebral fractures, with minimal overlap of reports amongst different terms, apart from 'Loss of Height' and 'Compression Fractures'. 'Loss of Height' was by far most effective term when compared to both 'Compression Fracture' and 'Crush Fracture'. In the present study, 'Loss of Height' had a greater probability of detecting a vertebral fracture compared to the other terms, however, the wide confidence interval in the odds ratios suggests the sample size was possibly too small.

During the Implementation phase, only 20, or less than a third of patients with vertebral fractures were invited to the SFP program with a large number of patients already on appropriate therapy. The latter may be due to the presence of a well-established and effective SFP program at Concord Hospital, which is a reflection of the excellent ability of the service to identify patients with fragility fractures. The median age of patients identified with vertebral fractures in the Implementation phase of the study was 79.6 years, which is consistent with the expected increase in vertebral fractures with ageing. However, the SFP program is not always accessible to frail, elderly patients who often reside in nursing homes. Therefore, this highlights the central important of primary care physicians in helping to manage these patients who could potentially be alerted regarding the presence of a vertebral fracture.

Strengths and Limitations of our Study

The strength of this study lies in its simple search strategy. It builds a bridge between previous studies (19, 20, 22, 25) and the application of NLP methods to the current clinical setting using existing readily available software and hardware. The Development phase identified the key search terms 'Loss of Height', 'Compression Fracture' and 'Crush Fracture' using a sample time-frame of six months. These terms can be easily implemented by any other SFP or health care service to identify patients with fragility fractures. With the added benefit of more user friendly software, our methods can be easily implemented on a regular basis for capturing this unique subset of patients.

The main limitation of our study was the search software, Centricity RIS-I 4.2 Plus. This is a pre-existing software in Concord Hospital that is poorly designed and not user friendly in relation to data interrogation and data extraction. The software itself allows limited flexibility in terms of limiting search criteria such as date range, nor did it enable the use of multiple search terms simultaneously. Further, it only allowed output to a text file, for manual review. During the Implementation phase, we were thus required to run searches using each term on a specified date, rather than being able to run the search term using a pre-specified date range. As a consequence, the process of identifying patients with vertebral fractures was time consuming and therefore difficult to systematise and thus apply to a clinical setting on a regular basis.

Unlike other NLP studies (26), the approach we used was comparatively simple and did not require dictionaries of coded medical terms for named entity recognition, nor complex negation detection algorithms. The average clinician is able to run a similar search to obtain their data without the need for vast amounts of knowledge in the field of natural language processing and software coding. Grundmeier et al (19) identified a number of studies that have used similar methods to extract relevant data to high degrees of accuracy and sensitivity. Sevenster et al (21) extracted measurement information with extremely high accuracy using text pattern matching techniques (recall and precision both >0.99).

Our study highlights that simple text based search systems are valuable in mining pools of free-text medical records. We feel that it should contribute to the growing literature supporting the use of NLP to help facilitate the adaptation of such technology integration into our current health care system.

Natural Language Processing – Is there a Role for it?

Natural language processing has the ability to mine free-text radiology reports, in order to populate clinically rich databases from which we can readily extract information to improve health care. It is an effective tool to allow quick mining of data from within free-text EMRs. NLP methods allow for flexibility; from simple text based searches to searches with more rigorous criteria. Once documents have been processed, different approaches and search strategies to identify specific outcomes can be implemented at a relatively low programming effort (23).

NLP methods are easily replicable and transferrable to another institution using the same or similar free-text search software without the needed for advanced knowledge in software coding. Further, they are far more scalable and replicable than manual data extraction, potentially allowing surveillance of entire health care system populations rather than a sub-sample (21, 23). Additionally, with the advancement of EMRs, there exists the opportunity to expand these methods to an outpatient setting. Murff et al. comments on the possibility of using search strategies on progress notes and investigation reports on a daily basis to prospectively monitor patients, thus giving the added benefit of identifying complications whilst a patient is still in the hospital, giving a more real-time quality assurance process (23).

Despite its many benefits, there are a number of limitations to using NLP methods for data extraction. These include, the heavy reliance on an up-to-date database from which to extract the data required. Missing reports, reports not uploaded correctly or reports that pre-date the EMR era would not be included in the search database.

The English language itself proves a problem when using NLP methods to mine free-text data such as radiology reports for useful information. A paper by Nadkarni et al. highlights in detail the importance of grammar, word/phrase order variation, derivation, inflection, synonymy and homographs, all which add to the difficulty in extracting meaningful accurate data (27). Further, the importance of negation and uncertainty identification, that is inferring whether a named entity is present or absent requires the use of a more comprehensible NLP method for accurate data extraction (27). Grundmeier RW et al, used simple methods to successfully extract data for long bone fractures from radiology reports, proving that there is value to their use. Further they state their sensitivity analysis performed similarly when compared to more complex methods (19).

The importance of accurate transcription of reports is utmost when it comes to utilising simple search strategies. This is because these text based methods allow little room for transcription errors. An error in typing the search term will result in that report not being identified using the search algorithm. Further to this, in a period of time constraints with large volumes of work, it is also important that a comment on the pathology in question has been made on the report. For example, a CT scan for abdominal pain may not necessarily warrant the reporting of an incidental finding of a fragility fracture of the spine when another more acute pathology such as bowel obstruction is present. Consequently, the ability of the search method to accurately identify correct reports also relies heavily on the transcription and initial reporting of all relevant findings.

The future of NLP methods and its integrating into the health care system seems promising. With the widespread role out of EMRs, NLP methods will play a vital role in improving the delivery of health care to the greater public. In its current form, NLP can put millions of clinical reports at the fingertips of researchers and clinicians (25). With the use of specifically designed software and search strategies reports can be mined for useful data with a relatively higher degree of accuracy.

The development of search strategies does require a knowledge of medicine and software, as well as several hours of effort to test their effectiveness. However, once established, they are easily replicable and applicable to similar studies across different centres (25). One must also be mindful of “concept drift”; changes in reporting styles and description of terminology as time passes. Thus the future application of NLP methods would require regular reviews and quality assurance checks in order to ensure they maintain their high sensitivity and positive predictive values.

CONCLUSION

We successfully created a simple search strategy based on natural language processing methods, and pre-existing hardware and software to mine free-text radiology reports to identify patients with fragility fractures. Our results indicate that the best search terms to identify patients with fragility fractures was ‘Loss of Height’, ‘Compression Fracture’ and ‘Crush Fracture’. These terms can be easily applied on a daily basis to search radiology reports and identify these important subset of patients. Our methods are simple and easily replicable, allowing them to be used at other sites where pre-existing search software may already exist. However, there is a need to apply the same technique to other healthcare facilities using more modern software in order to ensure its clinical application.

Our study highlights that simple text based search systems are valuable in mining pools of free-text medical records. We feel that it should contribute to the growing literature supporting the use of NLP to help facilitate the adaptation of such technology integration into our current health care system.

REFERENCES

1. Organization WH. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO study group [meeting held in Rome from 22 to 25 June 1992]. 1994.

2. Bliuc D, Nguyen ND, Nguyen TV, Eisman JA, Center JR. Compound risk of high mortality following osteoporotic fracture and refracture in elderly women and men. Journal of Bone and Mineral Research. 2013;28(11):2317-24.

3. Warrs J, Abimanyi-Ochom J, Sanders K. Osteoporosis costing Australians: A New Burden of Disease Analysis 2012 to 2022. Osteoporosis Australia. 2013.

4. Kanis JA, Johnell O, Oden A, Sernbo I, Redlund-Johnell I, Dawson A, et al. Long-Term Risk of Osteoporotic Fracture in Malmö. Osteoporosis International. 2000;11(8):669-74.

5. Nguyen ND, Ahlborg HG, Center JR, Eisman JA, Nguyen TV. Residual lifetime risk of fractures in women and men. Journal of Bone and Mineral Research. 2007;22(6):781-8.

6. Ahmed LA, Center JR, Bjørnerem Å, Bluic D, Joakimsen RM, Jørgensen L, et al. Progressively increasing fracture risk with advancing age after initial incident fragility fracture: the Tromsø study. Journal of Bone and Mineral Research. 2013;28(10):2214-21.

7. Center JR, Bliuc D, Nguyen TV, Eisman JA. RIsk of subsequent fracture after low-trauma fracture in men and women. JAMA. 2007;297(4):387-94.

8. Johnell O, Kanis J, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Fracture risk following an osteoporotic fracture. Osteoporosis international. 2004;15(3):175-9.

9. Melton Iii L, Atkinson E, Cooper C, O'Fallon W, Riggs B. Vertebral fractures predict subsequent fractures. Osteoporosis International. 1999;10(3):214-21.

10. Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA. Mortality after all major types of osteoporotic fracture in men and women: an observational study. The Lancet. 1999;353(9156):878-82.

11. Black DM, Arden NK, Palermo L, Pearson J, Cummings SR. Prevalent vertebral deformities predict hip fractures and new vertebral deformities but not wrist fractures. Journal of bone and mineral research. 1999;14(5):821-8.

12. Ismail AA, O'Neill TW, Cooper C, Finn JD, Bhalla AK, Cannata JB, et al. Mortality Associated with Vertebral Deformity in Men and Women: Results from the European Prospective Osteoporosis Study (EPOS). Osteoporosis International. 1998;8(3):291-7.

13. Black DM, Delmas PD, Eastell R, Reid IR, Boonen S, Cauley JA, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. New England Journal of Medicine. 2007;356(18):1809-22.

14. Cranney A, Guyatt G, Griffith L, Wells G, Tugwell P, Rosen C. IX: Summary of meta-analyses of therapies for postmenopausal osteoporosis. Endocrine reviews. 2002;23(4):570-8.

15. Cranney A, Papaioannou A, Zytaruk N, Hanley D, Adachi J, Goltzman D, et al. Parathyroid hormone for the treatment of osteoporosis: a systematic review. Canadian Medical Association Journal. 2006;175(1):52-9.

16. Papapoulos S, Chapurlat R, Libanati C, Brandi ML, Brown JP, Czerwiński E, et al. Five years of denosumab exposure in women with postmenopausal osteoporosis: results from the first two years of the FREEDOM extension. Journal of Bone and Mineral Research. 2012;27(3):694-701.

17. Andrade SE, Majumdar SR, Chan KA, Buist DS, Go AS, Goodman M, et al. Low frequency of treatment of osteoporosis among postmenopausal women following a fracture. Archives of Internal Medicine. 2003;163(17):2052-7.

18. Elliot-Gibson V, Bogoch ER, Jamal SA, Beaton DE. Practice patterns in the diagnosis and treatment of osteoporosis after a fragility fracture: a systematic review. Osteoporosis International. 2004;15(10):767-78.

19. Grundmeier R, Masino A, Casper T, Dean J, Bell J, Enriquez R, et al. Identification of Long Bone Fractures in Radiology Reports Using Natural Language

Processing to support Healthcare Quality Improvement. Applied Clinical Informatics. 2016;7(4):1051-68.

20. Lacson R, Khorasani R. Practical examples of natural language processing in radiology. Journal of the American College of Radiology. 2011;8(12):872-4.

21. Sevenster M, Buurman J, Liu P, Peters J, Chang P. Natural Language Processing Techniques for Extracting and Categorizing Finding Measurements in Narrative Radiology Reports. Applied clinical informatics. 2015;6(3):600-10.

22. Do BH, Wu AS, Maley J, Biswal S. Automatic Retrieval of Bone Fracture Knowledge Using Natural Language Processing. Journal of Digital Imaging. 2013;26(4):709-13.

23. Murff HJ, FitzHenry F, Matheny ME, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. JAMA. 2011;306(8):848-55.

24. Lih A, Nandapalan H, Kim M, Yap C, Lee P, Ganda K. Targeted intervention reduces refracture rates in pateitns with incident non-vertebral osteoporotic fractures: a 4-year prospective controlled study. Osteoporosis International. 2011;22(3):849-58.

25. Hripcsak G, Austin JHM, Alderson PO, Friedman C. Use of Natural Language Processing to Translate Clinical Information from a Database of 889,921 Chest Radiographic Reports. Radiology. 2002;224(1):157-63.

26. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. Journal of Biomedical Informatics. 2001;34(5):301-10.

27. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. Journal of the American Medical Informatics Association. 2011;18(5):544-51.